Wild Birds Use an Ordering Rule to Decode Novel Call Sequences

Suzuki, Toshitaka N ; Wheatcroft, David ; Griesser, Michael

DOI: https://doi.org/10.1016/j.cub.2017.06.031

Wild birds use an ordering rule to decode novel call sequences

Toshitaka N. Suzuki[1,2,6,*], David Wheatcroft[3], Michael Griesser[4,5]

[1]Center for Ecological Research, Kyoto University, 2-509-3 Hirano, Otsu, Shiga 520-2113, Japan

[2]Department of Evolutionary Studies of Biosystems, Hayama, Kanagawa 240-0193, SOKENDAI (The Graduate University for Advanced Studies), Japan

[3]Department of Animal Ecology, Uppsala University, Norbyvägen 18D, 75236 Uppsala, Sweden

[4]Department of Anthropology, University of Zurich, Winterthurerstrasse 190, 8057 Zürich, Switzerland

[5]Institute of Environmental Sciences, Jagiellonian University, Gronostajowa 7, 30-387 Krakow, Poland

[6]Lead Contact

[*]Correspondence: toshi.n.suzuki@gmail.com

SUMMARY

The generative power of human language depends on grammatical rules, such as word ordering, that allow us to produce and comprehend even novel combinations of words [1–3]. Several species of birds and mammals produce sequences of calls [4–6], and, like words in human sentences, their order may influence receiver responses [7]. However, it is unknown whether animals use call ordering to extract meaning from truly novel sequences. Here, we use a novel experimental approach to test this in a wild bird species, the Japanese tit (*Parus minor*). Japanese tits are attracted to mob a predator when they hear conspecific alert and recruitment calls ordered as alert-recruitment sequences [7]. They also approach in response to recruitment calls of heterospecific individuals in mixed-species flocks [8, 9]. Using experimental playbacks, we assess their responses to artificial sequences in which their own alert calls are combined into different orderings with heterospecific recruitment calls. We find that Japanese tits respond similarly to mixed-species alert-recruitment call sequences and to their own alert-recruitment sequences. Importantly, however, tits rarely respond to mixed-species sequences in which the call order is reversed. Thus, Japanese tits extract a compound meaning from novel call sequences using an ordering rule. These results demonstrate a new parallel between animal communication systems and human language, opening new avenues for exploring the evolution of ordering rules and compositionality in animal vocal sequences.

Keywords: bird; communication; compositionality; language; Paridae; syntax

RESULTS AND DISCUSSION

Japanese tits produce combinations of calls in particular contexts, which evoke specific responses in receivers [7]. They produce alert (ABC) calls when detecting a variety of predatory threats, recruitment (D) calls when attracting flock members in non-dangerous situations, and they combine these two calls into alert-recruitment (ABC-D) sequences when recruiting other flock members to mob a stationary predator (Figure 1A; [7, 10]). Previous playback experiments showed that receivers respond to alert (ABC) calls by scanning the surroundings for a potential danger, recruitment (D) calls by approaching the sound source, and alert-recruitment (ABC-D) sequences by mixing these two responses, i.e., scanning for danger while simultaneously approaching the sound source [7]. However, when the call order is artificially reversed (D-ABC), tits rarely exhibit either alert or approaching responses [7].

These findings suggest that tits use call ordering (i.e., alert-recruitment ordering rule) to extract a compound meaning from call sequences, which provides one of the first examples of compositional expressions in non-human animals [3, 11, 12]. However, Japanese tits may not assess the meanings of individual call units in a sequence, but rather perceive the whole sequence as an idiomatic expression with a unique message (e.g., "mobbing" instead of "alert + approach"; see [13–15]). Furthermore, it is also possible that they may respond more strongly to naturally combined sequences (ABC-D) simply because these sequences are more familiar than artificially reversed ones (D-ABC). If Japanese tits derive compound meanings from combinations of calls only when the meanings are given in a specific order, we predict that they should also derive compound meanings from novel combinations following the same ordering pattern (e.g., alert-recruitment ordering).

Here, we explore whether Japanese tits use an ordering rule to decode novel call sequences. Japanese tits form mixed-species flocks with willow tits (*Poecile montanus*) during the non-breeding season. Willow tits produce "tää" calls to attract both conspecific and heterospecific flock members, including Japanese tits (Figure 1B; [8, 9]). Like D calls of Japanese tits, "tää" calls are used in a variety of contexts, such as when arriving at a new foraging patch or when mobbing a predator [16–18]. Thus, despite differences in their acoustic structure, both willow tit's tää calls and Japanese tit's D calls have a general recruitment function. We use this synonymy to artificially generate novel call sequences by

combining conspecific alert calls (ABC) and heterospecific recruitment calls (tää) into different orderings (ABC-tää and tää-ABC sequences; Figure 1C). If tits use an alert-recruitment ordering rule to decode novel call sequences, they are expected to respond to artificial ABC-tää sequences similarly to synonymous ABC-D sequences, but only weakly to reversed tää-ABC sequences. In contrast, if they do not use a call ordering rule when hearing the mixed call sequences, they are expected to show similar responses to ABC-tää and tää-ABC sequences.

We tested the response of free-living Japanese tits (n = 28) to the playback of ABC-tää or tää-ABC call sequences. Each stimulus contained 30 call sequences in 90 seconds, matching the structure of natural calling bouts in predator mobbing contexts [10]. After locating a flock of Japanese tits, call sequences were played back from a loudspeaker hung from a tree branch. Behavioral observations were made on a focal individual that was the closest individual to the loudspeaker among the flock members when each trial was started. All the playback trials were conducted at different locations separated by at least 400 m, ensuring that all the data were collected from different individuals [8] (see STAR Methods).

Responses to Novel Sequences Depend on Call Ordering

During the playback of ABC-tää sequences, Japanese tits scanned the surroundings by turning their heads from side-to-side (median = 13 head turns per 90 s), while at the same time approaching the loudspeaker with hops and flights to within 2 m (n = 12/14 focal birds; Figure 2). Seven out of 14 focal individuals approached to within 1 m of the loudspeaker. In contrast, in response to tää-ABC sequences, tits made fewer head turns scans (median = 5 per 90 s; least-square mean: $Z = 4.47$, $p < 0.001$) and more rarely approached the loudspeaker within 2 m (n = 2/14 focal birds; $Z = 3.32$, $p < 0.001$; Figure 2). Also, only 1 out of 14 focal individuals approached within 1 m. Thus, when hearing novel combinations of ABC and tää calls, Japanese tits exhibit stronger responses when the component calls follow the alert-recruitment ordering rule.

The responses of Japanese tits to ABC-tää sequences were indistinguishable to those expressed in response to ABC-D sequences (Figure 2; scans: $Z = 0.20$, $p = 0.84$; approach within 2 m:

$Z = 1.29$, $p = 0.20$; see also [7]), although we note that ABC-D playbacks were conducted in a previous year, not allowing us to account for potential year-to-year variation. Together, these results indicate that, when hearing novel ABC-tää sequences, Japanese tits extract information equivalent to their own ABC-D sequences and that, like sequences containing ABC and D calls, call ordering determines responses.

ABC-first Ordering Does Not Explain Responses

Although Japanese tits discriminate between ABC-tää and tää-ABC sequences, it might be possible that they respond to any sequence beginning with ABC, irrespective of the call that follows, because their native and, therefore, familiar mobbing call sequences (ABC-D) typically begin with this call type. We assessed this possibility by combining alert calls of Japanese tits (ABC calls) with alert calls of willow tits (zi calls; Figure 1B; [18–20]), producing ABC-zi and zi-ABC sequences (Figure 1C). If ABC calls appearing first in a sequence is enough to elicit scanning and approach behavior, tits should respond more to ABC-zi than to zi-ABC sequences. In contrast, if ABC appearing first is not sufficient to explain responses to a given call sequence, tits are expected to respond similarly to ABC-zi and zi-ABC sequences. We tested the responses of Japanese tits ($n = 28$) to the playbacks of these two sequences in the same manner as the former experiment (see STAR Methods).

In response to both ABC-zi and zi-ABC sequences, Japanese tits exhibited some horizontal scans (ABC-zi: median = 4; zi-ABC: median = 5), but rarely approached the loudspeaker (within 2m: ABC-zi: n = 1/14 focal birds; zi-ABC: n = 2/14 focal birds; Figure 2). They produced similar numbers of scans ($Z = 0.46$, $p = 0.64$) and were similarly likely to approach ($Z = 0.60$, $p = 0.55$) in response to playbacks of either of these two sequences. Therefore, order-dependent responses to novel sequences depend on the specific calls included rather than whether or not ABC calls appear first.

Furthermore, our data allow us to reject the possibility that any modifier after ABC evokes mobbing-like responses in Japanese tits. Focal individuals exhibited fewer numbers of scanning and rarely approached the loudspeaker during the playback of ABC-zi sequences compared with the playback of ABC-tää sequences (scans: $Z = 2.52$, $p = 0.02$; approach: $Z = 3.38$, $p < 0.01$) and that of

ABC-D sequences (scans: $Z = 2.82$, $p = 0.01$; approach: $Z = 2.75$, $p < 0.01$). Thus, simply hearing ABC followed by a different call type is not sufficient to evoke either scanning or approach responses in Japanese tits.

Acoustic Similarity Does Not Drive Responses

Although D and tää calls differ in a number of acoustic features (Figure 1A), there remains a possibility that some features cause Japanese tits to perceive tää calls as their own D calls, and, consequently, perceive ABC-tää calls as ABC-D sequences (see [21]). To test this possibility, we compared the responses of Japanese tits to playbacks of D calls, tää calls, and artificially modified tää calls. The artificially modified tää calls were generated by shortening the duration of tää calls to 50% of their natural length, making their duration similar to that of D calls (Figure S1A) and to the calls of other related species within the genus *Poecile* [22]. A spectrographic cross-correlation analysis confirmed that modified short-tää calls have greater similarity to D calls than original tää calls (Figure S1B). If tits perceive D and tää calls as the same call type because of their acoustic similarity, then they are expected to approach in response to short-tää calls in a similar way to D calls and original tää calls. In contrast, if Japanese tits categorize D and tää calls as distinct calls albeit with similar meanings, they are expected to approach D and tää calls, but not the unfamiliar short-tää calls. We tested the response of Japanese tits ($n = 42$ flocks) to each of the three call types (60 calls in 180 seconds).

Japanese tits approached the loudspeaker during the playbacks of D calls and tää calls. A larger proportion of flock members approached within 2 m of the loudspeaker during the playback of D calls (median = 83% of flock members) than during the playback of tää calls (median = 13%) ($Z = 3.49$, $p < 0.001$). However, tits almost never approached the playback of short-tää calls (median = 0%; Figure 3) (short-tää vs D calls: $Z = 4.01$, $p < 0.001$; short-tää vs tää calls: $Z = 2.29$, $p = 0.02$). These results indicate that Japanese tits categorize D and tää calls as different call types, but extract a similar "approach" meaning from these calls. Therefore, equivalent responses to ABC-tää and ABC-D sequences are evoked by similarity in the meanings of their component calls, but not by similarity in their acoustic features.

Compositionality Drives Responses to Call Sequences

Japanese tits respond to the playbacks of novel alert-recruitment call sequences (ABC-tää) and their own alert-recruitment sequences (ABC-D) by simultaneously scanning and approaching, behaviors that tits express independently when hearing alert (ABC) and recruitment (tää and D) calls, respectively (Figure 2; [7]). The fact that Japanese tits fail to respond to novel short-tää calls suggests that they extract information from willow tit tää calls only after learning to associate them with recruitment contexts [23], rather than recognizing any calls with similar acoustic structure innately [21]. Weak responses to ABC-zi sequences allow us to reject the possibility that responses to ABC-D and ABC-tää sequences are driven by a process where ABC calls prime tits to expect and, therefore, perceive any following call as a recruitment call (Figure 2). Moreover, given that tits perceive tää and D calls as different calls (Figure 3), equivalent responses to ABC-tää and ABC-D sequences are not caused by similarity in acoustic features of component calls. Instead, our results strongly support the idea that tits extract an equivalent, compound meaning from both their own (ABC-D) and novel call sequences (ABC-tää) by assessing and combining the meanings of individual call units ("alert + approach"), but not by perceiving the whole sequences as an idiomatic expression (e.g., unique "mobbing" message).

Compositionality Depends on Call Ordering

Japanese tits exhibit mixed alert and approach responses to novel alert-recruitment sequences (ABC-tää), but only weakly respond to the reversed recruitment-alert sequences (tää-ABC). This result demonstrates that, like when responding to their own sequences (ABC-D), Japanese tits use call ordering (i.e., alert-recruitment) to decode novel sequences. These findings contrast with those from a previous study on monkeys. Although Diana monkeys (*Cercopithecus diana*) respond to heterospecific Campbell's monkey (*Cercopithecus campbelli*) "boom" calls when combined with Campbell's alarm calls, they ignore boom calls when artificially combined with conspecific alarm calls [24]. Diana monkeys do not produce boom calls or an equivalent call themselves [5, 24], and thus, are unlikely to

have an own ordering rule for the recognition of boom-related call combinations. As a result, Diana monkeys may not understand the semantic link between Campbell's boom calls and their own alarm calls. In contrast, Japanese tits use an ordering rule when combining their own calls into sequences (ABC-D sequences) [10] and extract similar information from willow tit's tää calls and their own D calls, which may allow them to generalize their ordering rule to novel ABC-tää sequences.

A critical aspect of compositionality in human language is that the meanings of word sequences (e.g., "John's small, blue socks") are not always generated from simply combining the meanings of the component words (e.g., "John", "small", "blue", "socks"). In this case, grammatical rules, such as word order, help us understand the intended meaning of the overall expression; such modification of meanings is known as non-trivial compositionality [12]. In contrast, trivial compositionality corresponds to combinations where the meaning of a sequence (e.g., "small, blue") is generated from a simple concatenation of the component words ("small" + "blue") [12]. Our experiments show that Japanese tits produce compound responses to sequences only when the calls are ordered into alert-recruitment sequences with a natural inter-call interval (i.e., 0.1-s) [7], suggesting that they might extract meaning in a non-trivial way. However, it is also possible that pragmatic rules (i.e., ordering of information units), but not grammatical rules, help tits discriminate between different orderings of call sequences [25, 26]. Future studies with artificially manipulated inter-call intervals could help to determine the cognitive mechanisms underlying the perception of call sequences.

Evolutionary Implications and Future Direction

It remains challenging to assess the evolutionary drivers that facilitate the evolution of call ordering rules. Clearly, both vocal signals and call ordering should be under strong selection to ensure that the message is transmitted effectively to receivers [11, 27]. In natural ABC-D call sequences of Japanese tits, the more urgent alert call is usually given before less urgent social recruitment call. Urgency first call sequences have been found in other birds [28, 29] and mammals ([30], but see [31]), allowing signalers to focus the attention of receivers on the predatory danger and, likewise, benefiting receivers by allowing them to modulate their behavior adaptively. In Japanese tits, alert calls unambiguously

communicate a risky situation and always appear at the beginning of sequences, while the subsequent recruitment calls are given in a range of different contexts that vary in their risk [7, 10]. Similarly, in call sequences of Campbell's monkeys, calls with less specific information are produced after more informative calls, where suffixation of "oo" generalizes the meaning of predator-specific alarm calls [32, 33]. No matter the reason, once a specific call ordering becomes fixed, it is likely that receivers would express weaker responses to aberrant orderings.

To conclude, our findings demonstrate that Japanese tits use an ordering rule to extract meanings from novel call sequences, representing a new analogy between animal communication systems and human language [3, 12]. Japanese tits and other members of the family Paridae are known to produce multiple note types and show a corresponding diversity of call combinations [22, 34]. The function of this vocal elaboration is still poorly understood, but the results presented here suggest that it may be underlain by call ordering rules that allow individuals to increase the number of messages that they can generate by combining their calls and associated meanings. Uncovering and comparing animal vocal sequences across species will likely reveal new aspects of the sophistication and complexity in animal communication systems, which may shed light on the evolution of compositionality and grammatical rules in language.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
    - Responses to novel call sequences
    - Comparisons between ABC-first sequences
    - Acoustic similarity between D and tää calls
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental information includes a figure and can be found with this article online at

AUTHOR CONTRIBUTIONS

T.N.S., D.W., M.G. designed research; T.N.S. performed research; T.N.S. and D.W. analyzed data; and T.N.S., D.W., M.G. wrote the paper. The authors declare no conflict of interest.

ACKNOWLEDGMENTS

We thank Carel van Schaik and four anonymous reviewers for valuable comments on the manuscript and Koppuno Fuchiko for support. This work was supported by JSPS KAKENHI Grant Number 16752305 (to TNS) and the National Science Centre, Poland, through the European Union's Horizon

2020 research and innovation programme (Marie Sklodowska-Curie grant No. 665778) (to MG).

REFERENCES

1. Chomsky, N. (1965). Aspects of the Theory of Syntax. (MIT Press, Cambridge, MA).

2. Fitch, W.T. (2010). Evolution of Language. (Cambridge Univ. Press, Cambridge, UK).

3. Hurford, J.R. (2011). The Origins of Grammar: Language in the Light of Evolution II. (Oxford University Press, Oxford).

4. Collier, K., Bickel, B., van Schaik, C.P., Manser, M.B., and Townsend, S.W. (2014). Language evolution: syntax before phonology? Proc. R. Soc. Lond. B *281*, 20140263.

5. Schlenker, P., Chemla, E., Schel, A. M., Fuller, J., Gautier, J.-P., Kuhn, J., Veselinovic, D., Arnold, K., Cäsar, C., Keenan, S., et. al. (2016). Formal monkey linguistics. Theor. Ling. *42*, 1–90.

6. Suzuki, T.N. (2016). Semantic communication in birds: evidence from field research over the past two decades. Ecol. Res. *31*, 307–319.

7. Suzuki, T.N., Wheatcroft, D., and Griesser, M. (2016). Experimental evidence for compositional syntax in bird calls. Nat. Commun. *7*, 10986.

8. Suzuki, T.N. (2012). Long-distance calling by the willow tit, *Poecile montanus*, facilitates formation of mixed-species foraging flocks. Ethology *118*, 10–16.

9. Suzuki, T. N. & Kutsukake, N. (2017). Foraging intention affects whether willow tits call to attract members of mixed-species flocks. R. Soc. Open Sci. *4*, 170222.

10. Suzuki, T.N. (2014). Communication about predator type by a bird using discrete, graded and combinatorial variation in alarm calls. Anim. Behav. *87*, 59–65.

11. Marler, P. (1998). Animal communication and human language. In The Origin and Diversification of Language, N. G. Jablonski, and L. C. Aiello, eds. (Univ. California Press), pp. 1–20.

12. Hurford, J. R. (2014). The Origins of Language: A Slim Guide. (Oxford University Press, Oxford).

13. Arnold, K., and Zuberbühler, K. (2006). Language evolution: Semantic combinations in primate calls. Nature *441*, 303.

14. Arnold, K., and Zuberbühler, K. (2008). Meaningful call combinations in a non-human primate. Curr. Biol. *18*, R202–R203.

15. Arnold, K., Zuberbühler, K. (2012). Call combinations in monkeys: Compositional or idiomatic expressions? Brain Lang. *120*, 303–309.

16. Suzuki, T.N. (2012). Calling at a food source: Context-dependent variation in note composition of combinatorial calls in willow tits. Ornithol. Sci. *11*, 103–107.

17. Haftorn, S. (1993). Ontogeny of the vocal repertoire in the willow tit *Parus montanus*. Ornis Scand. *24*, 267–289.

18. Haftorn, S. (2000). Contexts and possible functions of alarm calling in the Willow Tit, *Parus montanus*; the principle of 'better safe than sorry'. Behaviour *137*, 437–449.

19. Rajala, M., Rätti, O., and Suhonen, J. (2003). Age differences in the response of willow tits (*Parus montanus*) to conspecific alarm calls. Ethology *109*, 501–509.

20. Kareksela, S., Härmä, O., Lindstedt, C., Siitari, H., and Suhonen, J. (2013). Effect of willow tit *Poecile montanus* alarm calls on attack rates by Pygmy Owls *Glaucidium passerinum*. Ibis *155*, 407–412.

21. Randler, C. (2012). A possible phylogenetically conserved urgency response of great tits (*Parus major*) towards allopatric mobbing calls. Behav. Ecol. Sociobiol. *66*, 675–681.

22. Lucas, J.R., and Freeberg, T.M. (2007). “Information” and the chick-a-dee call: Communicating with a complex vocal system. In Ecology and Behavior of Chickadees and Titmice: An Integrated Approach, K.A. Otter, ed. (Oxford University Press, Oxford), pp. 199–213.

23. Magrath, R.D., Haff, T.M., McLachlan, J.R., and Igic, B. (2015). Wild birds learn to eavesdrop on heterospecific alarm calls. Curr. Biol. *25*, 2047–2050.

24. Zuberbühler, K. (2002). A syntactic rule in forest monkey communication. Anim. Behav. *63*, 293–299.

25. Schlenker, P., Chemla, E., and Zuberbühler, K. (2016). What do monkey calls mean? Trends Cogn. Sci. *20*, 894–904.

26. Schlenker, P., Chemla, E., Schel, A. M., Fuller, J., Gautier, J.-P., Kuhn, J., Veselinovic, D., Arnold, K., Cäsar, C., Keenan, S., et. al. (2016). Formal monkey linguistics: The debate. Theor. Ling. *42*, 173–201.

27. Searcy, W. A., and Nowicki, S. (2005). The Evolution of Animal Communication: Reliability and Deception in Signaling Systems. (Princeton University Press, Princeton, NJ).

28. Engesser, S., Ridley, A.R., and Townsend, S.W. (2016). Meaningful call combinations and compositional processing in the southern pied babbler. Proc. Natl. Acad. Sci. USA *113*, 5976–5981.

29. Freeberg, T. M. (2008). Complexity in the chick-a-dee call of Carolina chickadees (*Poecile carolinensis*): Associations of context and signaler behavior to call structure. Auk *125*, 896–907.

30. Swan, D. C. and Hare, J. F. (2008). The first cut is the deepest: Primary syllables of Richardson's ground squirrel, *Spermophilus richardsonii*, repeated call alert receivers. Anim. Behav. *76*, 47–54.

31. Arnold, K., and Zuberbühler, K. (2006). The alarm-calling system of adult male putty-nosed monkeys, *Ceropithecus nictitans martini*. *72*, 643–653.

32. Ouattara, K., Lemasson, A., and Zuberbühler, K. (2009). Campbell's monkeys use affixation to alter call meaning. PLoS ONE *4*, e7808.

33. Ouattara, K., Lemasson, A., and Zuberbühler, K. (2009). Campbell's monkeys concatenate vocalizations into context-specific call sequences. Proc. Natl. Acad. Sci. USA *106*, 22026–22031.

34. Krams, I., Krama, T., Freeberg, T.M., Kullberg, C., and Lucas, J.R. (2012). Linking social complexity and vocal complexity: A parid perspective. Philos. Trans. R. Soc. Lond. B *367*, 1879–1891.

35. Suzuki, T.N. (2011). Parental alarm calls warn nestlings about different predatory threats. Curr. Biol. *21*, R15–R16.

36. Kroodsma, D.E., Byers, B.E., Goodale, E., Johnson, S., and Liu, W.C. (2001). Pseudoreplication in playback experiments, revisited a decade later. Anim. Behav. *61*, 1029–1033.

37. Charif, R. A., Waack, A. M., and Strickman, L. M. (2010). Raven Pro 1.4. User's Manual, Cornell Lab of Ornithology, Cornell University. Available at: http://www.birds.cornell.edu/

38. R Core Team (2014). R: a language and environment for statistical computing (R Foundation for Statistical Computing). Available at: http://www.r-project.org/

39. Oksanenm J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P. et al. (2017). vegan: community ecology package. R package version 2.4-3. Available at: https://cran.r-project.org/web/packages/vegan/index.html

40. Ripley, B. (2015). MASS: support functions and datasets for Venables and Ripley's MASS. R package version 7.3-40. Available at https://cran.r-project.org/web/packages/MASS/index.html.

41. Lenth, R.V. (2015). lsmeans: least-squares means. R package version 2.17. Available at https://cran.r-project.org/web/packages/lsmeans/index.html.

42. Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. Roy. Stat. Soc. B *57*, 289–300.

Figure Legends

Figure 1. Experimental Setup

(A) Sound spectrograms of the natural alert-recruitment call sequence of Japanese tits given when mobbing a predator.

(B) Sound spectrograms of separate recruitment and alert calls of willow tits.

(C) Sound spectrograms of the novel call sequences composed of calls of Japanese and willow tits. Similarly to the natural call sequences of Japanese tit, calls from two species are artificially combined with an interval of 0.1 s.

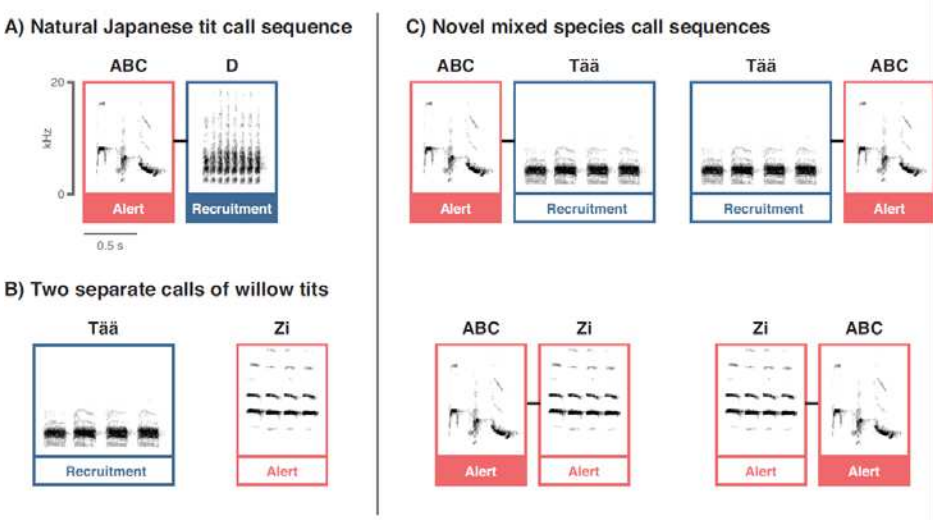


Figure 2. Response of Japanese Tits to Playbacks of Natural and Novel Call Sequences

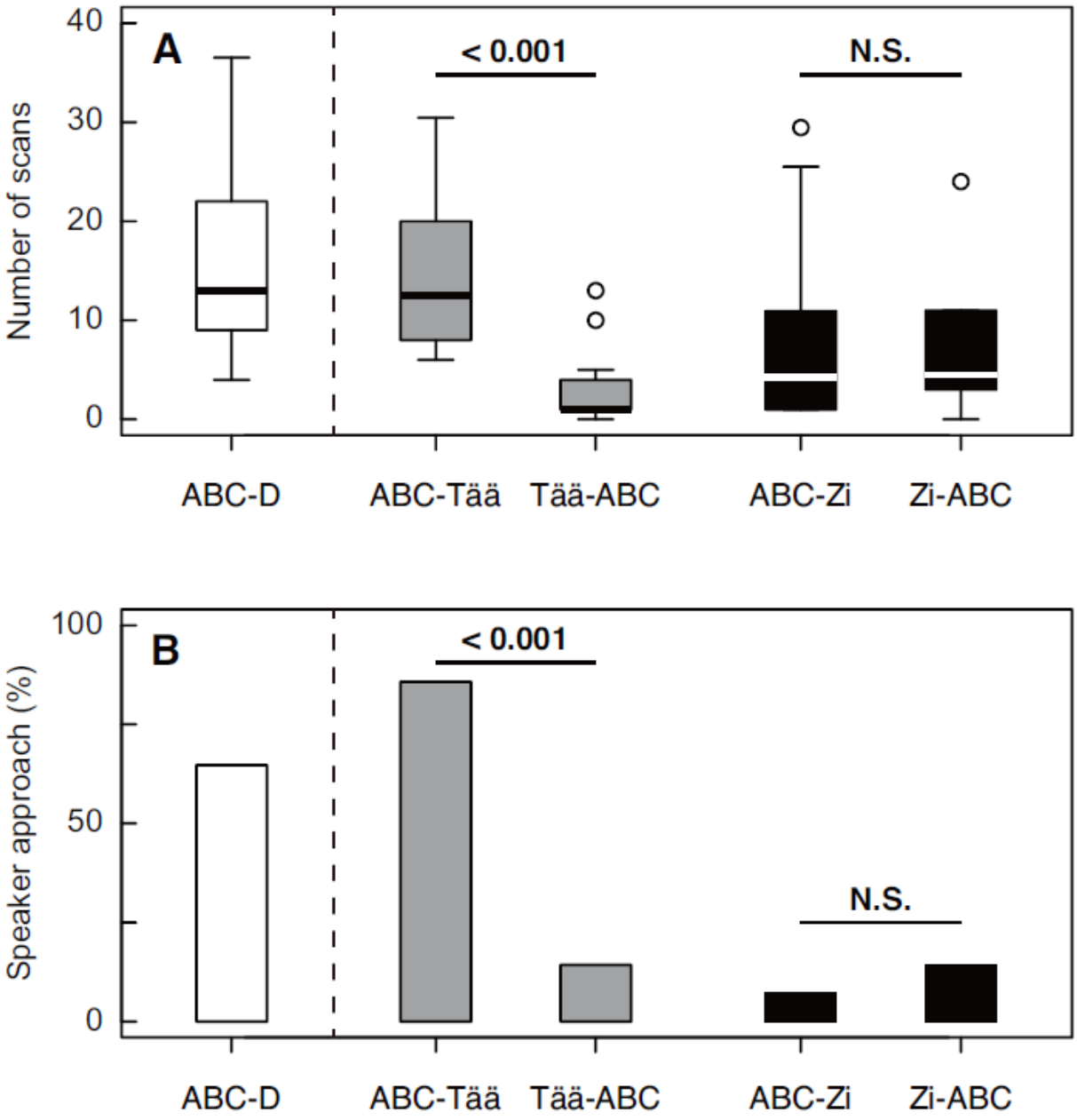
(A) Number of horizontal scans made by the focal individuals during 90-s. The box and whisker plots display the median and 25 and 75% quartiles; the whiskers are extended to the most extreme value inside the 1.5-fold interquartile range.

(B) Percentage of trials in which the focal individuals approached within 2 m of the loudspeaker.

(A–B) Natural call sequences (ABC-D) consists of their own calls only, whereas novel call sequences

348 are composed of their own alert calls (ABC) and either recruitment (tää) or alert (zi) calls of willow tits.
349 Data for ABC-D are derived from Experiment 2 of our previous study [7]. Sample size: n = 17
350 individuals for ABC-D; n = 14 individuals for each of ABC-tää, tää-ABC, ABC-zi, and zi-ABC. Each
351 individual was exposed to only one treatment.

352

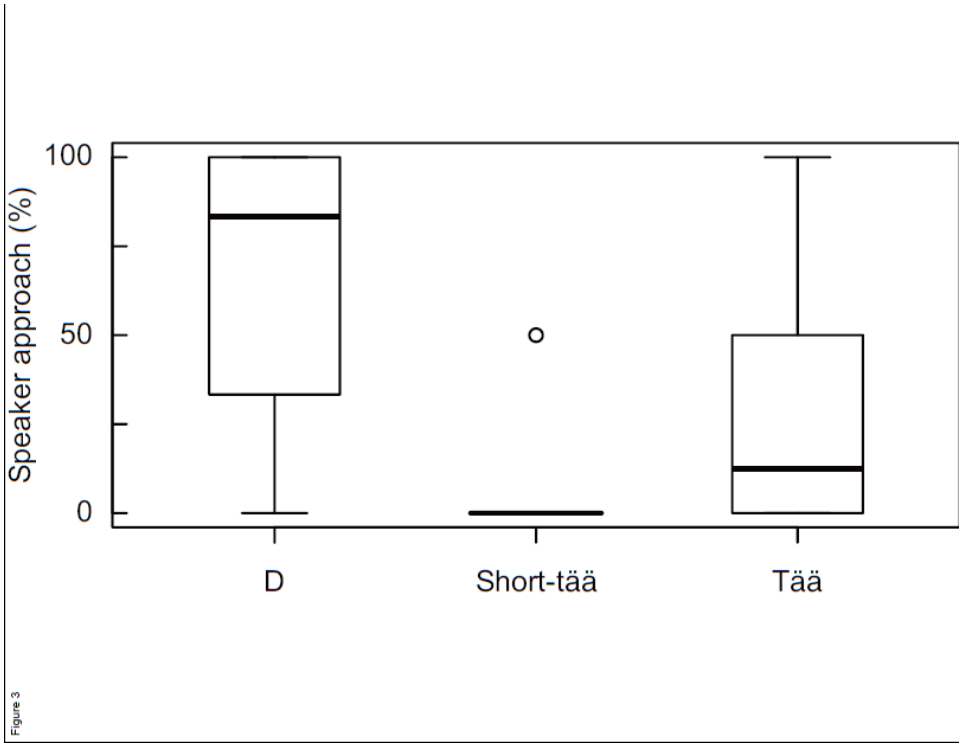


353

354

Figure 3. Responses to D and Tää Calls by Japanese Tits Are Not Caused by Sound Similarity.

Proportion of flock members that approached within 2 m of the calls during 3-min of playbacks. Japanese tits frequently approached within 2 m of the loudspeaker during playback of D calls and often during tää calls. However, they rarely approached in response to artificially shortened tää calls having an increased similarity to D calls (2m: $\chi^2 = 41.8$, $df = 2$, $p < 0.0001$). Sample size: n = 14 flocks for each treatment. See also Figure S1.



STAR★METHODS

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Toshitaka N. Suzuki (toshi.n.suzuki@gmail.com).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Data were collected from wild Japanese tits (56 individuals for the first experiment, 42 flocks for the second experiment). Experiments were performed in accordance with relevant guidelines and regulations. All the experimental protocols were approved by SOKENDAI Ethics Committee for Animal Research and adhered to the Guidelines for the Use of Animals in Research of the Animal Behavior Society/Association for the Study of Animal Behaviour. This research was performed under permission from the Ministry of the Environment and the Forestry Agency of Japan.

METHOD DETAILS

Responses to novel call sequences

We designed experiments to test whether Japanese tits use an alert-recruitment call ordering rule to extract information from novel call sequences. If this is the case, tits are expected to show stronger responses to novel call sequences that follow the alert-recruitment ordering than to those do not. To test this hypothesis, we prepared two novel sequences by combining the alert call of Japanese tits (ABC) and the recruitment call of willow tits (tää) into different orderings; ABC-tää and tää-ABC sequences (Figure 1C). However, because Japanese tits always produce ABC calls before D calls when combining them, it is possible that familiarity with ABC-first sequences causes any differential response. To test this possibility, we prepared two additional sequences by combining the alert call of Japanese tits (ABC) and that of willow tits (zi) to produce ABC-zi and reversed zi-ABC sequences (Figure 1C). Both tää and

zi calls of willow tits are often produced in mixed-species flocks [8, 9, 16], and thus, Japanese tits are familiar with both of these calls.

The calls used for this study were recorded during previous field experiments by TNS [10, 35]. ABC notes of Japanese tits were recorded by exposing either a stuffed jungle crow (*Corvus macrorhynchos*) or a stuffed Japanese marten (*Martes melampus*) near their nest boxes in May and June 2010. Zi and tää calls of willow tits were recorded by exposing a stuffed sparrowhawk (*Accipiter nisus*) near artificial feeders from December 2008 to January 2009. All the source individuals for the playback calls were individually identified either by unique combinations of color-rings or by their nest boxes. Calls were recorded with an LS370 parabolic microphone (Fuji Planning Corporation, Tokyo, Japan) connected to an R-09HR digital audio recorder (sampling wave files at 48kHz and 16bits; Roland Corporation, Shizuoka, Japan) or to an MZ-RH1 Hi-MD walkman (sampling wave files at 44.1 kHz and 16bits; Sony Corporation, Tokyo, Japan) at the same field site as this study.

We created playback stimuli by using Adobe Audition 3.0 software. We chose three types of notes (A, B, and C) from our recording library of Japanese tits (n = 14 individuals) and two types of notes (zi and tää) from the recording library of willow tits (n = 14 individuals for both) on the basis of the sound quality. A, B, and C notes were then combined into an ABC sequence with natural intervals between the notes (50–150 ms, measured for each individual of the record source). Zi notes and tää notes were typically produced as a string of several notes (zi: 1-10 notes per call; tää: 1-20 notes per call). Thus, we chose calls with four zi notes and those with four tää notes as the acoustic modifiers of ABC.

To generate playback stimuli, we combined ABC calls with either zi or tää calls into ABC-tää, tää-ABC, ABC-zi, and zi-ABC sequences. Each call in a given sequence was separated at 0.1 s, which is within the range of intervals of ABC and D notes of Japanese tits [7]. Then, each call sequence was repeated in a sound file at a rate of 20 call sequences per minute (one sequence every 3 s, total duration 90 s). This calling rate corresponds to the range of the natural repetition rate of mobbing call sequences of Japanese tits [10]. Within each playback stimulus, each call sequence was separated from any preceding sequences by background noise of either 1.1 ± 0.1 s (ABC-tää and tää-ABC sequences)

or 1.7 ± 0.1 s (ABC-zi and zi-ABC sequences) (mean ± s.d., n = 14 for each). This difference in intervals reflects the differences in the duration of tää and zi calls (tää: 1.2 ± 0.1 s, zi: 0.7 ± 0.0 s; n = 14 for each). Fourteen unique playback stimuli were created for each of the four treatments. We saved the sound files in WAV format (16-bit accuracy, 44.1-kHz sampling rate) onto an SD memory card.

We conducted playback experiments in mixed deciduous and coniferous forests near Karuizawa, Nagano Prefecture, Japan (36°17'–23'N, 138°28'–39'E). First, we located a flock of Japanese tits. Upon finding a flock, we hung an AT-SPG50 loudspeaker (Audio-Technica Corporation, Tokyo, Japan) from a tree branch at 1.8 ± 0.1 m (mean ± s.d., n = 56) from the ground. The loudspeaker was connected to an R-09 HR digital audio recorder (Roland Corporation, Shizuoka, Japan) with extension cords, which enabled the control of playbacks from an observation position *ca.* 10 m away from the loudspeaker. We chose the individual that was closest to the loudspeaker as the focal individual and recorded its behavior during the playback. We played back calls at a standardized volume (75 dB re 20 mPa at 1 m from the loudspeaker measured using an SM-325 sound level meter; AS ONE Corporation, Osaka, Japan), which is similar to the natural amplitude of calls given by Japanese tits and willow tits.

To determine the tits' responses to different treatments, we recorded the following behavioral variables during 90-s of playbacks: (1) number of horizontal scans: we counted the number of movements that focal birds made with their heads from left to right or right to left (approximately a 180° turn) and (2) approaching the loudspeaker: we recorded whether focal birds approached within 2 m of the loudspeaker during the playback. The observations were made with binoculars, and these variables were commented on an R-09 HR digital audio recorder. For birds which flew to a tree with the loudspeaker, we measured the minimum distance from the bird to the loudspeaker after the trials.

Trials were carried out under calm and dry weather conditions between 08:45 and 15:30 h (Japan Standard Time), from 20 October to 2 November 2016, when Japanese tits often form mixed-species flocks with willow tits. Treatments with the same call composition (e.g., ABC-tää and tää-ABC) were alternated in successive trials, and two-trial blocks with different call composition (i.e., [ABC-tää and tää-ABC] or [ABC-zi and zi-ABC]) were played back in turns, so that responses to all four treatments were observed under largely similar conditions. In two trials, a black kite (*Milvus migrans*)

or a jungle crow flew close to the focal bird, and tits were apparently alerted to these potential threats. In these cases, we stopped the trials and repeated the same treatment in the next place. The minimum distance between experimental sites was 400 m, because previous observations of color-banded individuals showed that this distance is enough to ensure that independent data are collected from different individuals [8].

We collected data from 42 males and 14 females across all four treatments (ABC-tää: 11 males and 3 females; tää-ABC: 12 males and 2 females; ABC-zi: 11 males and 3 females; zi-ABC: 8 males and 6 females). The sex of the birds was determined from their sexually dimorphic plumage patterns (male Japanese tits have a broader black stripe from the breast to the vent than females). Each focal individual was exposed to a unique exemplar of each call sequence to avoid pseudo-replication [36].

Comparisons between ABC-first sequences

We investigated whether tits derive an equivalent information ("approach with alertness") from ABC-tää sequences and synonymous ABC-D sequences, but not from ABC-zi sequences. In 2015, we conducted playback experiments with ABC-D sequences in the same manner as the present experiment: we created a total of n = 17 unique exemplars of ABC-D sequences that contain 30 calls in 90-s files (one call every 3 s), playing them to Japanese tits in the wild (n = 17; 11 males and 6 females) [7]. Using the data from this previous experiment, we compared the responses of tits to three sequences; ABC-D, ABC-tää, and ABC-zi sequences. We note that we could not control for the possible effect of the study year.

Acoustic similarity between D and tää calls

The present experiment is based on the assumption that Japanese tits perceive their own D calls and heterospecific tää calls as different call types, but extract similar information from them. However, it is still possible that tits may simply respond to the acoustic features shared between these two calls (see

[21]). To assess this possibility, we carried out a playback experiment using D calls, tää calls, and novel type of "tää" calls having an increased similarity to D calls. Using the software program Audacity 2.1.0 (http://audacity.sourceforge.net/), we artificially shortened the duration of tää calls (1.23 ± 0.08 s; mean ± s.d., n = 14) to be 50% (0.62 ± 0.04 s, n = 14), producing "short-tää" calls (Figure S1A). These short-tää calls have a similar duration to that of D calls (0.67 ± 0.09 s, n = 14), while keeping the same frequency-related measures of the original tää calls. If tits respond to D and tää calls simply because of their acoustic similarity, tits are expected to approach to the novel short-tää calls similarly or more strongly than to tää calls. In contrast, if some other mechanism, such as learning [23], explains responses to tää calls, tits should have reduced responses to novel, unfamiliar short-tää calls compared to original tää calls. We obtained tää calls (n = 14) comprised of four tää notes from different individual willow tits and edited them to obtain short-tää calls (n = 14). We also prepared D calls comprised of 7-10 D notes recorded from different individual Japanese tits (n = 14). Each of the sound files contains each type of call repeated at a rate of 20 calls per minute (one call every 3 s, total duration 180 s).

To confirm whether we successfully increased the similarity of tää calls by simply shortening the duration, we carried out a spectrographic cross-correlation (SPCC) analysis using the batch correlation function in Raven Pro 1.5.0 [37]. SPCC calculates sound similarity by comparing two sound spectrograms at all possible time points to produce a correlation coefficient [37]. All the calls were bandpass filtered at 1000-16000 Hz to reduce the possible influence of background noise in analyses. Spectrograms were input using a 3 dB filter bandwidth of 248 Hz, with a Hann window function (50% overlap of time grid and 256 samples of frequency grid). Then, we obtained two matrices of correlation coefficients from the comparisons between D calls and tää calls and between D calls and short-tää calls. In these matrices, each value varies between 0 to 1; greater values represent larger similarity to D calls. These matrices were then compared by using a Mantel test with 5000 permutations (Figure S1B). This statistical test was conducted using R for Mac OS X version 3.1.2 [38] and its function (*mantel* in the R package *vegan*; [39]).

We conducted this playback experiment in mixed deciduous and coniferous forests near

Karuizawa, Nagano Prefecture, Japan (36°18'–22'N, 138°31'–39'E). Similarly to the first experiment, we hung the loudspeaker from a tree branch (height: 1.7 ± 0.1 m, mean ± s.d., n = 42) at a distance of *ca.* 20 m from a flock of Japanese tits. The loudspeaker was connected to an R-09 HR digital audio recorder (Roland Corporation, Shizuoka, Japan) with extension cords, which enabled the control of playbacks from an observation position *ca.* 10 m away from the loudspeaker. We counted the number of flock members and then played back calls at a standardized volume (75 dB re 20 mPa at 1 m from the loudspeaker). During the 3-min of call playback, we recorded the number of Japanese tits approached within 2 m of the loudspeaker. The observations were made with binoculars, and these variables were commented on an R-09 HR digital audio recorder.

Trials were carried out under calm and dry weather conditions between 08:45 and 15:30 h (Japan Standard Time), from 25 March to 30 March 2017 when Japanese tits form mixed-species flocks with willow tits. The order of treatments was counter-balanced, so that responses to all three treatments were observed under largely similar conditions. In two trials, a jungle crow and a sparrowhawk flew close to the focal flock and tits were apparently alerted to them. In these cases, we stopped the trials and repeated the same treatment in the next place. The minimum distance between experimental sites was 400 m, to ensure the collection of independent data from different individual tits [8]. We collected data from 42 flocks across all three treatments (14 flocks per each of the three treatments). Each focal flock was exposed to a unique exemplar of each call to avoid pseudo-replication [36].

QUANTIFICATION AND STATISTICAL ANALYSIS

We analyzed the data using generalized linear models in R [38]. We used a negative binomial error distribution and log-link function (*glm.nb* in the R package *MASS*; [40]) for models analyzing the number of horizontal scans and a binomial error distribution and logit-link function (*glm* in the R package *stats*) for models analyzing the probability of approach behavior (yes or no).

First, we analyzed whether tits respond differently to the four types of playbacks (ABC- tää, tää-ABC, ABC-zi, zi-ABC). We fitted call ordering (ABC first or last), call composition (ABC + tää or

ABC + zi), and their interaction term, and sex of the focal bird as fixed terms. In 12 out of 56 trials, focal individuals flew into a dense canopy during the playbacks. Thus, we determined the time duration in which we could record on the behavior of the focal individual as the observation time and included this term in the analysis of horizontal scans as a log-transformed offset. In these analyses, we found a significant interaction between ordering and composition (scans: $\chi^2 = 7.64$, $df = 1$, $p < 0.01$; approach: $\chi^2 = 7.34$, $df = 1$, $p < 0.01$). However, we did not find any influence of sex of the focal birds on either behavioral response (scans: $\chi^2 = 0.13$, $df = 1$, $p = 0.72$; approach: $\chi^2 = 1.05$, $df = 1$, $p = 0.31$). Therefore, we further conducted post-hoc pairwise comparisons of responses between ABC-tää and tää-ABC calls, and ABC-zi and zi-ABC calls by using least-square means (*lsmeans* in the R package *lsmeans*; [41]), where sex was excluded from the models. We used likelihood ratio tests to calculate the significance of each fixed term.

For comparisons between ABC-tää, ABC-zi, and ABC-D sequences, we fitted playback treatments and sex as fixed terms. In 11 out of 31 trials (ABC-tää: 4/14; ABC-zi: 6/14; ABC-D: 7/17), tits flew out of sight during the observation. Thus, we included the time duration in which we could record the behavior of the focal individual and included this term in the analysis of horizontal scans as a log-transformed offset. In the primary analyses, call treatment had a significant effect on both responses (scans: $\chi^2 = 7.84$, $df = 2$, $p = 0.02$; speaker approach: $\chi^2 = 21.42$, $df = 2$, $p < 0.0001$), whereas sex had no significant influences on either response (scans: $\chi^2 = 2.17$, $df = 1$, $p = 0.14$; speaker approach: $\chi^2 = 0.08$, $df = 1$, $p = 0.77$). Therefore, we further compared the responses of tits between the three call treatments by using least-square means (*lsmeans* in the R package *lsmeans*; [41]), where sex was excluded from the models.

Data from playbacks of D, tää, and short-tää calls were analyzed using generalized linear models (*glm* in the R package *stats*) with a binomial error structure and logit link function for primary analysis. Responses of tits to the three call treatments were then compared by using least-square means (*lsmeans* in the R package *lsmeans*; [41]). We fitted the proportion of flock members approached within 2 m of the loudspeaker as a dependent variable and call treatment as a fixed term.

We used log likelihood ratio tests for calculating p-values. When making multiple

549 comparisons, a false discovery rate control [42] was used to adjust p-values. All tests were two-tailed
550 and the significance level was set at $\alpha = 0.05$.

551

552 **DATA AND SOFTWARE AVAILABILITY**

553 All software required to perform the analyses described in the 'Quantification and statistical analysis'
554 section is freely available to download for the open source R program. Raw data have been deposited in
555 Mendeley Data, http://dx.doi.org…